

## **РОЗРОБКА ТЕСТІВ ЗІ ЗМІНОЮ ФОРМИ ЗОБРАЖЕННЯ ОБ'ЄКТА УВАГИ ДЛЯ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ РОЗПІЗНАВАННЯ В ПРИКЛАДНИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМАХ**

Показано можливість створення тестів з шкалою зміни зображення форми об'єктів уваги для дослідження інтелектуальних систем розпізнавання візуальної інформації.

На даний час пропонується дуже багато різноманітних систем обробки і розпізнавання зображень різних класів для різних предметних областей. При цьому часто виникає потреба в'яснити: чи здатна дана система вирішувати конкретну задачу, чи можна її адаптувати до неї, чи треба шукати іншу систему. Основним і найбільш простим шляхом є створення відповідних тестових зображень з заданими характеристиками, які б дали відповідь на поставлені питання.

Розробка тестових зображень, до яких відносять, наприклад, штрихові міри чи телевізійні таблиці, мають важливе значення у своїх предметних вузькоспеціалізованих областях. В якості загальноприйнятих тестів в сфері обробки зображень широко використовуються емпіричні зображення-тести, наприклад зображення «Лена». В психології, офтальмології, використовують свої спеціалізовані тестові зображення.

Проте в загальному, усі ці тести є візуальними і практично не мають ні аналітичного (розрахункової формули), ні кількісного (статистичного) вираження їхньої вартості. Іншими словами, вони не можуть забезпечити встановлення залежності між входом і виходом атестованої системи, виступаючи в експериментальних дослідженнях, в якості «аргументу» (вхідного сигналу), результатом яких є «значення-функція» – реакція або прийняте рішення. Переважно основною оцінкою результату є бінарна величина: дві поділки шкали: добре – погано, так – ні. Звичайно в кращому положенні є тести типу штрихових мір чи таблиця визначення гостроти зору. Їхні параметри можуть бути такі виражені кількісно. Для інших, можливим є використання методів експертного оцінювання.

Тому побудова тестових зображень, розпізнавальна складність або складність відбору потрібної інформації є актуальною задачею і має важливе значення для автоматичних та автоматизованих систем розпізнавання, зокрема, стосовно визначення їхньої ефективності та можливостей вирішення як конкретних задач, на які вони запроектовані, так і різних інших супутніх задач, близьких або подібних в сенсі обробки та розпізнавання.

Метою даної роботи є розробка підходу до побудови тестів, на яких конфігурація форми зображення об'єкта уваги змінюється у відповідності до

величини конкретного параметра, який визначає обсяг вилучення високочастотних спектральних складових шляхом перетворення Фур'є.

**Особливості тестових зображень.** В реальних ситуаціях змістом розпізнавання є ідентифікація виявленого об'єкта з об'єктами конкретних класів. Тому, в зображеннях-тестах використовують один і той самий фон, на якому локалізують потрібні об'єкти різного розміру, орієнтації, у відповідних ракурсах бачення їхніх реальних обрисів або проекцій. Очевидно, що таких варіантів може бути безліч, а тому серед цих зображень-тестів відбирають лише декілька, найбільш в тому чи іншому сенсі прийнятних і найбільш адекватних даних задачі. Такий відбір переважно здійснюється експертними методами, причому в якості експертів виступають самі ж розробники тестів.

Зображення об'єктів на реальних сценах через втрати в каналах сприйняття, передачі сигналів та в процесі формування на виході візуального образу цієї сцени, отримують деяке спотворення, що призводить до хибної ідентифікації та зниження оперативності щодо прийняття рішення. В цьому плані, одним з аспектів тестування можна вважати і результати розпізнавання за різних рівнів спотворення. Якщо зображення сцени є кольоровим або має велику кількість градацій сірого, тобто є чорно-білим, то спотворенню можуть підлягати відтінки, контраст, чіткість. У випадку бінарних зображень, які використовують лише два кольори, переважно чорний і білий, один з яких належить тлу, а інший об'єкту спотворення стосуються лише конфігурації форми контуру зображення об'єкта уваги, і можуть бути як незначними, так і повністю виключити можливість правильної їх ідентифікації.

В тестових зображеннях такі спотворення мають свій позитивний сенс. Дійсно, чим за більших спотворень контуру об'єкта уваги система може його правильно розпізнати тим ефективнішою є робота цієї системи. В цьому плані, тестові зображення повинні забезпечити деякий набір об'єктів уваги з різним ступенем спотворення форми. Для цього необхідно забезпечити такі вимоги: по-перше, ступені спотворення форми повинні відповідати градаціям деякої шкали і по-друге, рівень спотворень повинен визначатися деяким параметром, який однозначно визначає даний рівень спотворення, є повторюваним, достатньо повно інтерпретованим і практично незалежним від характеристик зображення об'єкта.

Реально, такі спотворення зображення об'єкта мають місце при втратах в каналі передачі, формування візуального образу та попередньої обробки і проявляються вони через розмиття конфігурації контуру, його згладжування, зміни контрасту, кольоропередачі тощо. Модельовання цього типу спотворень контуру легко здійснити застосовуючи пряме та обернене двомірне дискретне перетворення Фур'є у випадку бінарних зображень. Щодо кольорових та півтонових зображень то існують певні труднощі. Застосування обох перетворень Фур'є полягає в тому, що спочатку отримують спектральний розклад зображення, а далі з нього вилучають

високочастотні складові і за допомогою оберненого перетворення з решти складових формують фактично нове зображення. Це зображення, в залежності від кількості вилучених високочастотних буде або дуже близьким за формою до оригіналу або цілком від нього відмінним. Даний підхід використано в [1] для дослідження просторово-часового спектру зорових образів, які збережені короткочасною пам'яттю зорової системи людини. В якості тестів використано зображення натурної сцени – інтер'єру кімнати, з якого ступенево вилучені високочастотні спектральні складові. В [2] для оцінювання складності штрихових зображень, в якості яких використано китайські ієрогліфи, також визнають і розглядають їх спектральний склад.

**Програмна реалізація перетворення Фур'є.** В даній роботі для дослідження зміни конфігурації форми бінарного зображення об'єкта уваги так само використовується пряме і обернене перетворення Фур'є у формі програми-інтерфейсу. Реалізація дискретного перетворення Фур'є розрахована на розмір вікна  $256 \times 256$ , що цілком є достатнім, оскільки розміри об'єктів уваги є досить малими порівняно з розміром зображення-тесту. Саме перетворення здійснюється безпосередньо з використанням алгоритму розбиття на стрічки і стовпці. Для малих розмірів об'єктів уваги кількість операцій перетворення і час є прийнятними, а тому швидкі алгоритми для цього перетворення не використовуються.

Програма перетворення реалізована мовою C++ в середовищі CBuilder. В даному випадку мова C++ використана для пояснення алгоритму реалізації дискретного прямого і оберненого перетворення Фур'є за таких міркувань. Справа в тому, що середовище CBuilder є дуже близьким до логічних і мислених процесів людини, а тому реалізовані цією мовою алгоритми працюють практично аналогічно до міркувань людини. Крім того, важливе значення має те, що C++ є кросплатформенною мовою програмування, яка дозволяє переносити програмний код у відмінні від Microsoft Windows XP операційні системи, і вже там здійснювати компілювання. Розроблений в такий спосіб програмний продукт можна переносити на такі операційні системи як: Unix, Linux, Mac OS та інші. Дана програма є діючою, легко доповнюється додатковими функціями, проте в даній роботі вона використовується лише для практичної перевірки ідеології вирішення поставленої задачі. Особливістю даного варіанта програми є те, що вона використовує лише бінарні зображення, розмір сторін яких є кратний степені двійки.

**Опис інтерфейсу програми.** В результаті виклику і запуску програми маємо простий інтерфейс для роботи користувача, який включає чотири клавіші, два великих вікна для зображення-оригіналу та зображення-образу та дев'яти інформаційних полів, призначення яких пояснюється написами біля кожного з них. Крім самого перетворення, якщо є необхідність виділити

контур спотвореного зображення, розрахувати основні кількісні характеристики для зображення форми, як перетворюваного зображення об'єкта так і для перетвореного, програма також забезпечує ці функції. Для цього використовується спеціальний вказівник. Результати розрахунків відображаються в інформаційних полях, кількість яких вибирається дослідником і включається в інтерфейс, як це зображено на рис.1.

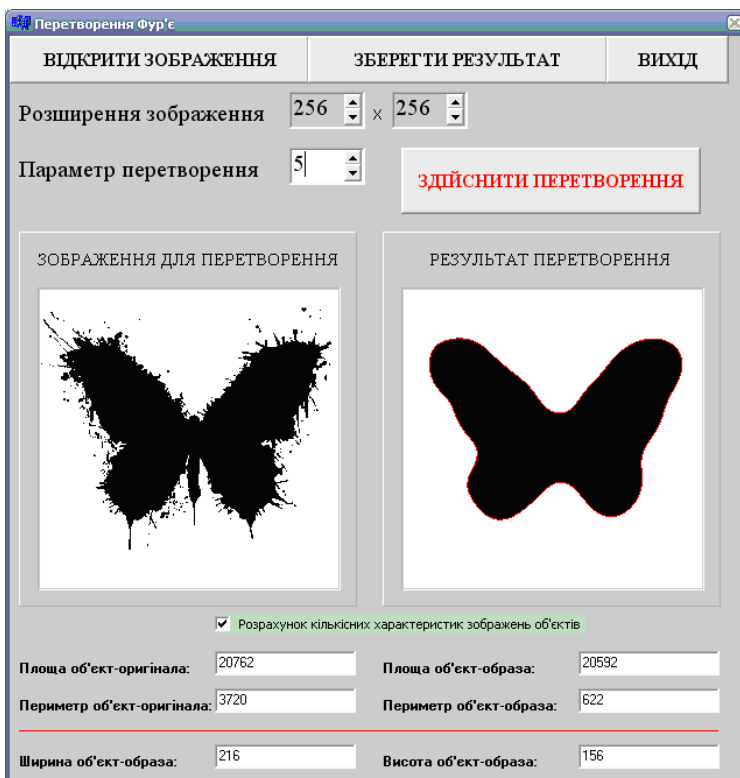


Рис.1. Зображення інтерфейсу перетворення Фур'є.

Для здійснення претворення необхідно ввести зображення або вирізаний відповідного розміру фрагмент з об'єктом, вказати значення параметра перетворення у відповідному інформаційному полі і запустити програму кнопкою «ЗДІЙСНИТИ ПЕРЕТВОРЕННЯ». Робота програми була перевірена для зображень з такими розмірами картинки: 64×64, 128×128, 192×192 та 256×256, для інших розмірів програма видає повідомлення про помилку. Результати перетворення можна зберігати, як на жорсткому диску ПК, так і на переносних носіях, для чого необхідно натиснути клавішу

“ЗБЕРЕГТИ РЕЗУЛЬТАТ” і вибрати вид збереження. Для завершення роботи і виходу з програми необхідно натиснути на клавішу “ВИХІД”.

**Дослідження впливу зменшення кількості гармонік.** В процесі дослідження впливу зменшення кількості гармонік на зміну конфігурації контуру, а відтак і на розпізнавальну властивість зображення було випробувано різні класи бінарних зображень об’єктів, два з яких зображено на рис.2. Враховуючи те, що на реальних зображеннях проблемних ситуацій практично об’єкти уваги переважно знаходяться серед інших об’єктів, які можуть відрізнятися розмірами, формою, орієнтацією, торкатись даних об’єктів або частково їх закривати, вибрані об’єкти містять незначний випадково локалізований точковий (імпульсний) шум.

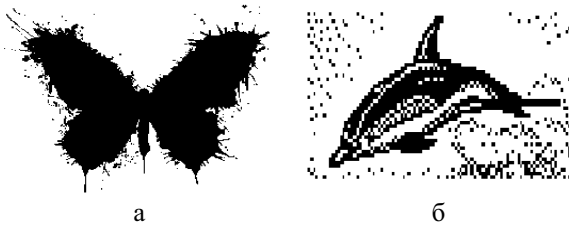


















Рис.2. Досліджувані об’єкти.

В даному експериментальному дослідженні використано зображення розміром 256x256 пікселів. Час перетворення становив порядку 9-12 секунд при тактовій частоті процесора 2.51 ГГц. При збільшенні розмірів зображення до величин, що є не кратними степені двійки та для зображень, в яких форма області покриття відмінна від квадратної, можуть виникати різні артефакти та спотворення, зумовлені в першу чергу малою кількістю складових та цілочисельністю елементів і параметрів перетворення за відсутності методів швидкого перетворення Фур’є. В табл.1 показано 16 зображень, причому для зображення об’єкта на рис.1а перетворювання здійснювали з кроком рівним  $\Delta s = 4$ , а для зображення об’єкта на рис.1б з кроком –  $\Delta s = 5$ . Для значень  $s = 30-22$  першого і значень  $s = 40-30$  другого зображень відмінності від оригіналу не дуже помітні і об’єкти розпізнаються правильно. Для значень  $s = 18-10$  першого і значень  $s = 25-15$  другого форма є суттєво спотвореною, проте розпізнавання ще може бути правильним. Для значень  $s = 6-2$  першого і значень  $s = 10-5$  другого форма практично не відповідає оригіналам і розпізнавання не може бути правильним.

На рисунках в таблиці зі зменшенням значення  $S$  поступово зникають шуми, стає більш гладким контур, який вироджується в плавні геометричні опуклі фігури. Зображення контуру стає нечітким і розпізнати його практично неможливо.

Таблиця 1

Зміна форми об'єкта уваги в залежності від величини параметра $S$			
			
30	26	22	18
			
14	10	6	2
			
40	35	30	25
			
20	15	10	5

Основними метрологічними величинами для більшості показників є площа і периметр пікселя. Величину площі зображення об'єкта визначають простим підрахунком кількості пікселів відмінних від фону за допомогою сканування усього зображення або лише області покриття об'єкта. Для визначення периметра підраховують кількість пікселів контурної лінії, отриманої в результаті виділення контуру. Цей спосіб не є коректним з кількох причин.

По-перше, контур може бути виділений, як за рахунок пікселів об'єкта, що прилягають до тла, так і за рахунок інверсування кольору пікселів тла, що прилягають до об'єкта.

По-друге, будь-якому контуру є властивий певний тип зв'язності, а отже, кількість пікселів в контурі буде різною.

По-третє, поняття периметра і контуру є різними. Контур це візуальне представлення форми бінарного об'єкта, отримане за допомогою замкненої послідовності пікселів, в той час, як периметр це метрична величина рівна сумі довжин сторін пікселів по лінії розділу, тобто по лінії «об'єкт – тло». Очевидно, що кількість пікселів контуру і кількість сторін пікселів буде різною, оскільки в залежності від положення пікселя, він може бути включений у периметр однією, двома, трьома або жодною із своїх сторін. Іншими словами, при визначенні периметра необхідно рахувати не кількість пікселів, а кількість сторін тих пікселів, які граничать з пікселями тла. Визначення величини периметра не складно реалізувати безпосередньо в процесі сканування зображення при виділенні контуру, використовуючи допоміжні відповідні предикати.

Низькочастотні складові сигналу несуть інформацію про локалізацію найбільш потужних складових спектру зображення об'єктів. Ці складові для впізнання об'єкта є найбільш важливі, оскільки вони в першу чергу привертають увагу, а тому мозок людини і виділяє їх.

Високочастотні складові зображення відповідають за різкі зміни в конфігурації форми контуру, тобто саме вони формують голкоподібні виступи і впадини на контурі. Іншими словами, чим більш зрізаною є форма конфігурації контуру, тим більший вклад мають високочастотні спектральні складові. Тому, усуваючи такі складові змінюється форма контуру в той спосіб, що він розмивається, знижується різкість деталей зображення.

**Висновки.** За результатами проведених дослідів можна стверджувати, що дійсно, деяка частина високочастотних складових спектру є надлишковими, тобто втрата деякої їхньої кількості є непомітною. Якщо зберегти тільки частину низькочастотних складових, то за рахунок цього відбудеться стиск зображення. Так працюють деякі алгоритми кодування.

При зниженні високочастотних гармонік зменшується інтенсивність шуму, але разом з тим і різкість самого об'єкта.

На підставі результатів проведених експериментів практичного дослідження перетворення Фур'є, можна стверджувати про можливість побудови шкали розпізнавальної складності зображень з різними градаціями спотворення конфігурації форми бінарного об'єкта, пов'язаної з значеннями параметра перетворення. Наведені в табл.1. зображення спотвореної форми відповідно до значення параметра перетворення представляють монотонну послідовність втрат високочастотних спектральних складових. Зазначимо також, що результати перетворень мають дискретний характер, зміст якого в тому, що і кількість складових в перетворенні і значення параметра перетворення є цілими числами. Для малорозмірних бінарних зображень об'єктів уваги це свідчить на користь створення невеликого набору зображень тестів, тим більше, що в даному випадку використовуються лише ті значення параметра перетворення, для яких спотворення є прийнятними.

З наведених в таблиці ілюстрацій, застосування безпосередньо прямого і оберненого перетворення Фур'є може бути використане для побудови шкали розпізнавальної складності зображень об'єктів уваги. Очевидно, що досить малі значення та відповідно великі значення параметра перетворення мають бути поза такою шкалою, оскільки в першому випадку спотворення форми є занадто сильно виражені, а в другому вони практично непомітні.

Тому, можна вважати, що запропонований підхід для розробки тестових зображень для оцінювання ефективності автоматичних та автоматизованих систем обробки та розпізнавання візуальної і визначення фізичного порогу рівня спотворення форми об'єкта уваги є цілком прийнятний.

1. Камінський Р.М. Градування складності розпізнавання зображень з використанням шумових полів // Праці Міжн. конф. з індуктивного моделювання. «МКІМ-2002». – Львів: ДНДІ Інформаційної інфраструктури, 2002. – Т.1, Ч.2. – С. 228 – 234.
2. Меркульєв А.В., Чихман В.Н., Шелепин Ю.Е. Исследование кратковременной памяти человека методами цифровой обработки изображений // Оптический журнал, том 68, № 5, 2001. – С. 36 – 42.
3. Бондарко В.М., Голузина А.Г., Данилова М.В., Чихман В.Н., Шелепин Ю.Е. Оценка сложности зрительных изображений // Сенсорные системы. – 2003. – том 17. – № 2. – С. 83 – 90.
4. Камінський Р.М. Властивості та характеристики малорозмірних об'єктів в тестових зображеннях систем розпізнавання та обробки зорових образів // Вісник Національного університету «Львівська політехніка» «Комп'ютерна інженерія та інформаційні технології». – 2003. – № 496. – С. 153 – 162.

*Поступила 15.09.2010р.*